Tourism Forecasting

Chiranjeevi Revanth Kumar Jarugula
Department of Computer Science
University of Bridgeport, Bridgeport, CT

Abstract

This paper is about the estimation of tourism demand which is a foundation of all tourism –related business decisions ultimately rest. Here forecasting of tourism for goa for goa state. The state tourism is the key success for largely business. The tourism forecasting developmental methods are of two categories they are quantitative and the qualitative methods. The quantitative forecasting is of 3 sub categories they are non casual time series model, time series model and artificial intelligent techniques. A time series model explains a variable with regard to it's own past and a random disturbance term. for the forecasting of the tourism data mining algorithms like linear regression is used and the multiplicative linear regression is also used it is widely used in the modeling and statistics. Here we forecast for the next 12 months tourism by using the past data. Here the data is available in the two forms they are of foreign tourists visiting goa and the domestic tourists visiting goa every month. Because of the tourism forecasting it uses for not only government but also other sectors like hotels, tourism providers etc.

Problem Statement

The objective of the forecasting is to enable Goa Government (and other stakeholders) to do forecasts for the next 12 months for state of Goa, month after month.

The data source for the analysis was the official website of department of tourism, Govt of Goa. We got monthly tourist visits from Jan 2009 to May 2015. The data was available in the form of two time series one for domestic tourist visiting Goa and other for foreign tourists visiting Goa . The domestic time series had an upward trend with yearly seasonality. The foreign time series did not have a trend but there was six month seasonality.

Forecasts and their assumptions –

We generated 12 months forecast in future along with their confidence intervals, i.e. the interval between which the forecast could vary. Some key assumptions for our forecasts are, Firstly data for at-least 12 months back is available for forecasting**,** secondly there won't be any huge macroeconomic changes in the world economy.

Conclusion

The final forecasting model recommended is the multiple linear regression model mentioned above. Secondly, we need to ensure that we have the latest data available while generating the forecast. This is based on the assumption that the govt. agencies and other stake holders preparing this forecast will have access to latest data which may not be published on the website. In case the data is not available then appropriate amount of error buffer should be built in while planning.

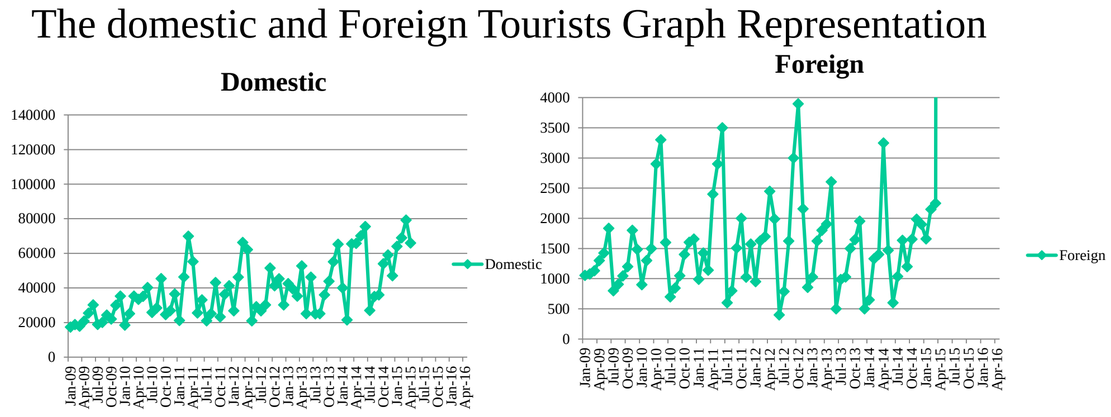
Data mining Algorithms And Evaluation Plan

Data Set

Here some of the data is printed

Month	Domestic	Foreign
January 2009	17,342	1054
February 2009	18,778	1078
March 2009	17,854	1132
April 2009	20,779	1300
May 2009	25,417	1425
June 2009	30,236	1835
July 2009	18,900	800
August 2009	20,025	900
September 2009	24,325	1045
October 2009	22,000	1200
November 2009	30,012	1800
December 2009	35,215	1485

The domestic and Foreign Tourists Graph Representation



Data Mining Algorithms

Model Description -

Multiple Linear Regression (MLR) for both domestic and foreign time-series is finalized, it is widely used in modeling and statistics for prediction. Here a version of multiplicative of this model i.e.

Demand = Fac1 * Fac2 * Fac3 * Fac4

Model Performance -

Here this model performance is much better than the naïve forecasts , i.e. previous K months accepting forecast as next months forecast. This value K was 6 in case of foreign naïve and 12 in case of domestic naïve.

Linear Regression

Linear Regression (Multiplicative)

Evaluation Plan

Linear Regression

- Here i carried out a linear regression of Demand Vs t, t^2 , lag12, monthly.
- Here tried different combinations, rejected this method, due to a very clear seasonality in residuals

Linear Regression (Multiplicative)

- Here regressed log(demand) Vs t, t^2, log(lag12), monthly.
- Here again tried different combinations, stuck to taking t, log(lag12) and monthly for domestic and t and monthly for foreign

The Predicted graphs for Domestic and Foreign Tourists are